



IFW

PATENT
81940.0075

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re application of:

KASAKO, et al.

Serial No: 10/820,629

Filed: April 8, 2004

For: Remote Storage Disk Control
Device With Function To Transfer
Commands To Remote Storage
Devices

Art Unit: 2186

Examiner: Not Assigned

I hereby certify that this correspondence is being deposited with the United States Postal Service with sufficient postage as first class mail in an envelope addressed to:

Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450, on
June 29, 2004

Date of Deposit

Shingale Ferguson

Name

Shingale Ferguson

Signature

June 29, 2004

Date

TRANSMITTAL OF PRIORITY DOCUMENT

Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

Dear Sir:

Enclosed herewith is a certified copy of Japanese patent application No. 2003-325082, which was filed September 17, 2003, from which priority is claimed under 35 U.S.C. § 119 and Rule 55.

Acknowledgment of the priority document(s) is respectfully requested to ensure that the subject information appears on the printed patent.

Respectfully submitted,

HOGAN & HARTSON L.L.P.

Date: June 29, 2004

By: *Anthony J. Orler*

Anthony J. Orler

Registration No. 41,232

Attorney for Applicant(s)

500 South Grand Avenue, Suite 1900
Los Angeles, California 90071
Telephone: 213-337-6700
Facsimile: 213-337-6701

日 本 国 特 許 庁
JAPAN PATENT OFFICE

別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office.

出 願 年 月 日 2 0 0 3 年 9 月 1 7 日
Date of Application:

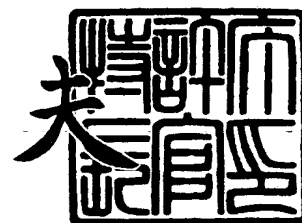
出 願 番 号 特 願 2 0 0 3 - 3 2 5 0 8 2
Application Number:
[ST. 10/C]: [J P 2 0 0 3 - 3 2 5 0 8 2]

出 願 人 株式会社日立製作所
Applicant(s):

2 0 0 4 年 3 月 2 4 日

特許庁長官
Commissioner,
Japan Patent Office

今 井 康 夫



出証番号 出証特 2 0 0 4 - 3 0 2 4 1 7 3

【書類名】 特許願
【整理番号】 340300154
【提出日】 平成15年 9月17日
【あて先】 特許庁長官殿
【国際特許分類】 G06F 3/06
【発明者】
 【住所又は居所】 神奈川県小田原市中里 3 2 2 番 2 号 株式会社日立製作所 R A I
 D システム事業部内
 【氏名】 加迫 尚久
【発明者】
 【住所又は居所】 神奈川県小田原市中里 3 2 2 番 2 号 株式会社日立製作所 R A I
 D システム事業部内
 【氏名】 近藤 修次
【発明者】
 【住所又は居所】 神奈川県小田原市中里 3 2 2 番 2 号 株式会社日立製作所 R A I
 D システム事業部内
 【氏名】 鈴木 亨
【特許出願人】
 【識別番号】 000005108
 【氏名又は名称】 株式会社日立製作所
【代理人】
 【識別番号】 110000176
 【氏名又は名称】 一色国際特許業務法人
 【代表者】 一色 健輔
【手数料の表示】
 【予納台帳番号】 211868
 【納付金額】 21,000円
【提出物件の目録】
 【物件名】 特許請求の範囲 1
 【物件名】 明細書 1
 【物件名】 図面 1
 【物件名】 要約書 1

【書類名】 特許請求の範囲**【請求項 1】**

情報処理装置と、
第一の記憶ボリュームを備える第一の記憶装置と、
第二の記憶ボリュームを備える第二の記憶装置と
を備え、
前記情報処理装置と前記第一の記憶装置とが通信可能に接続され、
前記第一の記憶装置と前記第二の記憶装置とが通信可能に接続され、
前記情報処理装置が、第一の通信規約に従い前記第一の記憶装置にデータの書き込みを
要求する第一の書き込み要求部を備え、
前記第一の記憶装置が、第二の通信規約に従い第二の記憶装置にデータの書き込みを要
求する第二の書き込み要求部を備える
記憶装置システムの制御方法において、
前記情報処理装置が、前記第二の記憶装置で実行される第一の命令を含んだ第一のデー
タを生成するステップと、
前記情報処理装置が、前記第一のデータを前記第一の通信規約に従い前記第一の記憶ボ
リュームに書き込む要求を、前記第一の書き込み要求部に送信するステップと、
前記第一の記憶装置が、前記第一の記憶ボリュームに書き込まれている前記第一のデー
タが前記第二の記憶装置に対する命令である場合に、前記第一のデータを前記第二の通信
規約に従い前記第二の記憶ボリュームに書き込む要求を、前記第二の書き込み要求部に送
信するステップと、
前記第二の記憶装置が、前記第二の記憶ボリュームに書き込まれている前記第一のデー
タに設定されている前記第一の命令を実行するステップと
を備えることを特徴とする記憶装置システムの制御方法。

【請求項 2】

前記情報処理装置が、前記第一の記憶装置で実行される第二の命令を含んだ第二のデー
タを生成するステップと、
前記情報処理装置が、前記第二のデータを前記第一の通信規約に従い前記第一の記憶ボ
リュームに書き込む要求を、前記第一の書き込み要求部に送信するステップと、
前記第一の記憶装置が、前記第一の記憶ボリュームに書き込まれている前記第二のデー
タが前記第一の記憶装置に対する命令である場合に、前記第二の命令を実行するステップ
と
を備えることを特徴とする請求項 1 に記載の記憶装置システムの制御方法。

【請求項 3】

前記第一の記憶装置が第三の記憶ボリュームを備え、
前記第二の記憶装置が第四の記憶ボリュームを備え、
前記第二の記憶装置が、第三の通信規約に従い第一の記憶装置にデータの読み取りを要
求する読み取り要求部を備える
記憶装置システムの制御方法において、
前記情報処理装置が、前記第三の記憶ボリュームが主記憶ボリューム、前記第四の記憶
ボリュームが副記憶ボリュームであるペアの形成命令を前記第一の命令として前記第一の
データを生成するステップと、
前記第二の記憶装置が、前記第二の記憶ボリュームに書き込まれている前記第一のデー
タに設定されている前記ペアの前記形成命令を受領し、前記第三の記憶ボリュームのデー
タを前記第三の通信規約に従い前記第四の記憶ボリュームに読み取る要求を、前記読み取
り要求部に送信するステップと
を備えることを特徴とする請求項 1 に記載の記憶装置システムの制御方法。

【請求項 4】

前記第一の記憶装置が第三の記憶ボリュームを備え、
前記第二の記憶装置が第四の記憶ボリュームと第六の記憶ボリュームとを備える

記憶装置システムの制御方法において、

前記情報処理装置が、前記第三の記憶ボリュームが副記憶ボリューム、前記第四の記憶ボリュームが主記憶ボリュームであるペアの形成命令を前記第一の命令として前記第一のデータを生成するステップと、

前記第二の記憶装置が、前記第二の記憶ボリュームに書き込まれている前記第一のデータに設定されている前記ペアの前記形成命令を受領し、前記第四の記憶ボリュームに書き込まれるデータの複製及び位置情報を前記第六の記憶ボリュームにジャーナルとして記憶するステップと

を備えることを特徴とする請求項 1 に記載の記憶装置システムの制御方法。

【請求項 5】

前記第一の記憶装置が第五の記憶ボリュームを備え、

前記第二の記憶装置が第六の記憶ボリュームを備え、

前記第三の記憶ボリュームが主記憶ボリューム、前記第四の記憶ボリュームが副記憶ボリュームであるペアが形成され、

前記第三の記憶ボリュームに書き込まれるデータの複製及び位置情報が前記第五の記憶ボリュームにジャーナルとして記憶されている

記憶装置システムの制御方法において、

前記情報処理装置が、前記ジャーナルの複製を前記第六の記憶ボリュームに記憶するジャーナル取得命令を前記第一の命令として前記第一のデータを生成するステップと、

前記第二の記憶装置が、前記第二の記憶ボリュームに書き込まれている前記第一のデータに設定されている前記ジャーナル取得命令を受領し、前記第五の記憶ボリュームのデータを前記第三の通信規約に従い前記第六の記憶ボリュームに読み取る要求を、前記読み取り要求部に送信するステップと

を備えることを特徴とする請求項 3 に記載の記憶装置システムの制御方法。

【請求項 6】

前記情報処理装置が、前記第六の記憶ボリュームに記憶されたジャーナルを用いて前記第四の記憶ボリュームを更新する復元命令を前記第一の命令として前記第一のデータを生成するステップと、

前記第二の記憶装置が、前記第二の記憶ボリュームに書き込まれている前記第一のデータに設定されている前記復元命令を受領し、前記第六の記憶ボリュームに書き込まれている前記ジャーナルに記憶されているデータを前記第四の記憶ボリュームに書き込むステップと

を備えることを特徴とする請求項 5 に記載の記憶装置システムの制御方法。

【請求項 7】

前記第二の記憶装置が、前記第三の記憶ボリュームを副記憶ボリューム、前記第四の記憶ボリュームを主記憶ボリュームとした前記ペアが形成されていることをペア管理情報として記憶するステップを備えることを特徴とする請求項 4 に記載の記憶装置システムの制御方法。

【請求項 8】

前記情報処理装置が、前記第三の記憶ボリュームを主記憶ボリューム、前記第四の記憶ボリュームを副記憶ボリュームとする前記ペアの交換命令を前記第一の命令として前記第一のデータを生成するステップと、

前記第二の記憶装置が、前記第二の記憶ボリュームに書き込まれている前記第一のデータに設定されている前記ペアの前記交換命令を受領し、前記ペア管理情報に、前記第三の記憶ボリュームを主記憶ボリューム、前記第四の記憶ボリュームとして記憶するステップと、

前記第二の記憶装置が、前記ペアの前記交換命令を受領し、前記第四の記憶ボリュームに書き込まれたデータの履歴を前記第六の記憶ボリュームにジャーナルとして記憶する処理を停止するステップと

を備えることを特徴とする請求項 7 に記載の記憶装置システムの制御方法。

【請求項 9】

情報処理装置と、
第一の記憶ボリュームを備える第一の記憶装置と、
第二の記憶ボリュームを備える第二の記憶装置と
を備え、
前記情報処理装置と前記第一の記憶装置とが通信可能に接続され、
前記第一の記憶装置と前記第二の記憶装置とが通信可能に接続され、
前記情報処理装置が、第一の通信規約に従い前記第一の記憶装置にデータの書き込みを
要求する第一の書き込み要求部を備え、
前記第一の記憶装置が、第二の通信規約に従い第二の記憶装置にデータの書き込みを要
求する第二の書き込み要求部を備える
記憶装置システムにおいて、
前記情報処理装置が、前記第二の記憶装置で実行される第一の命令を含んだ第一のデー
タを生成する命令設定部を備え、
前記情報処理装置が、前記第一のデータを前記第一の通信規約に従い前記第一の記憶ボ
リュームに書き込む要求を、前記第一の書き込み要求部に送信する命令送信部を備え、
前記第一の記憶装置が、前記第一の記憶ボリュームに書き込まれている前記第一のデー
タが前記第二の記憶装置に対する命令である場合に、前記第一のデータを前記第二の通信
規約に従い前記第二の記憶ボリュームに書き込む要求を、前記第二の書き込み要求部に送
信するデータ転送部を備え、
前記第二の記憶装置が、前記第二の書き込み部により前記第二の記憶ボリュームに書き
込まれている前記第一のデータに設定されている前記第一の命令を実行する第二の命令実
行部を備える
ことを特徴とする記憶装置システム。

【請求項 10】

前記情報処理装置が、前記第一の記憶装置で実行される第二の命令を第二のデータを生
成する前記命令設定部を備え、
前記情報処理装置が、前記第二のデータを前記第一の通信規約に従い前記第一の記憶ボ
リュームに書き込む要求を、前記第一の書き込み要求部に送信する前記命令送信部を備え
、
前記第一の記憶装置が、前記第一の記憶ボリュームに書き込まれている前記第二のデー
タが前記第一の記憶装置に対する命令である場合に、前記第二の命令を実行する第一の命
令実行部を備える
ことを特徴とする請求項 9 に記載の記憶装置システム。

【請求項 11】

前記第一の記憶装置が第三の記憶ボリュームを備え、
前記第二の記憶装置が第四の記憶ボリュームを備え、
前記第二の記憶装置が、第三の通信規約に従い第一の記憶装置にデータの読み取りを要
求する読み取り要求部を備える
記憶装置システムにおいて、
前記情報処理装置が、前記第三の記憶ボリュームが主記憶ボリューム、前記第四の記憶
ボリュームが副記憶ボリュームであるペアの形成命令を前記第一の命令として前記第一の
データを生成する前記命令設定部を備え、
前記第二の命令実行部が、前記第二の記憶ボリュームに書き込まれている前記第一のデー
タに設定されている前記ペアの前記形成命令を受領し、前記第三の記憶ボリュームのデー
タを前記第三の通信規約に従い前記第四の記憶ボリュームに読み取る要求を、前記読み
取り要求部に送信する形成コピー部を備える
ことを特徴とする請求項 9 に記載の記憶装置システム。

【請求項 12】

前記第一の記憶装置が第三の記憶ボリュームを備え、

前記第二の記憶装置が第四の記憶ボリュームと第六の記憶ボリュームとを備える記憶装置システムにおいて、

前記情報処理装置が、前記第三の記憶ボリュームが副記憶ボリューム、前記第四の記憶ボリュームが主記憶ボリュームであるペアの形成命令を前記第一の命令として前記第一のデータを生成する前記命令設定部を備え、

前記第二の記憶装置が、前記第二の記憶ボリュームに書き込まれている前記第一のデータに設定されている前記ペアの前記形成命令を受領し、前記第四の記憶ボリュームに書き込まれたデータの複製及び位置情報を前記第六の記憶ボリュームにジャーナルとして記憶するジャーナル記憶部を備える

ことを特徴とする請求項 9 に記載の記憶装置システム。

【請求項 13】

前記第一の記憶装置が第五の記憶ボリュームを備え、

前記第二の記憶装置が第六の記憶ボリュームを備え、

前記第三の記憶ボリュームが主記憶ボリューム、前記第四の記憶ボリュームが副記憶ボリュームであるペアが形成され、

前記第三の記憶ボリュームに書き込まれるデータの複製及び位置情報が前記第五の記憶ボリュームにジャーナルとして記憶されている

記憶装置システムにおいて、

前記情報処理装置が、前記ジャーナルの複製を前記第六の記憶ボリュームに記憶するジャーナル取得命令を前記第一の命令として前記第一のデータを生成する前記命令設定部を備え、

前記第二の命令実行部が、前記第二の記憶ボリュームに書き込まれている前記第一のデータに設定されている前記ジャーナル取得命令を受領し、前記第五の記憶ボリュームのデータを前記第三の通信規約に従い前記第六の記憶ボリュームに読み取る要求を、前記読み取り要求部に送信するジャーナル取得部を備える

ことを特徴とする請求項 11 に記載の記憶装置システム。

【請求項 14】

前記情報処理装置が、前記第六の記憶ボリュームに記憶されたジャーナルを用いて前記第四の記憶ボリュームを更新する復元命令を前記第一の命令として前記第一のデータを生成する前記命令設定部を備え、

前記第二の命令実行部が、前記第二の記憶ボリュームに書き込まれている前記第一のデータに設定されている前記復元命令を受領し、前記第六の記憶ボリュームに書き込まれている前記ジャーナルに記憶されているデータを前記第四の記憶ボリュームに書き込むリストア部を備える

ことを特徴とする請求項 13 に記載の記憶装置システム。

【請求項 15】

前記第二の命令実行部が、前記第三の記憶ボリュームを副記憶ボリューム、前記第四の記憶ボリュームを主記憶ボリュームとする前記ペアが形成されていることをペア管理情報として記憶するペア管理部を備えることを特徴とする請求項 12 に記載の記憶装置システム。

【請求項 16】

前記情報処理装置が、前記第三の記憶ボリュームを主記憶ボリューム、前記第四の記憶ボリュームを副記憶ボリュームとする前記ペアの交換命令を前記第一の命令として前記第一のデータを生成する前記命令設定部を備え、

前記第二の命令実行部が、前記第二の記憶ボリュームに書き込まれている前記第一のデータに設定されている前記ペアの前記交換命令を受領し、前記ペア管理情報に前記第三の記憶ボリュームを主記憶ボリューム、前記第四の記憶ボリュームとして記憶する前記ペア管理部を備え、

前記第二の命令実行部が、前記ペアの前記交換命令を受領し、前記第四の記憶ボリュームに書き込まれたデータの履歴を前記第六の記憶ボリュームにジャーナルとして記憶する

処理を停止するジャーナル停止部を備える
ことを特徴とする請求項 15 に記載の記憶装置システム。

【請求項 17】

第一の記憶ボリュームと、
通信可能に接続された第二の記憶装置に、第二の通信規約に従いデータの書き込みを要求する第二の書き込み要求部と
を備える記憶装置において、

前記第一の記憶ボリュームに書き込まれている第一のデータが前記第二の記憶装置に対する命令である場合に、前記第一のデータを前記第二の通信規約に従い前記第二の記憶装置が備える第二の記憶ボリュームに書き込む要求を、前記第二の書き込み要求部に送信するデータ転送部を備えることを特徴とする記憶装置。

【請求項 18】

第二の記憶ボリュームを備える記憶装置において、
前記第二の記憶ボリュームに書き込まれている第一のデータに設定されている第一の命令を実行する第二の命令実行部を備えることを特徴とする記憶装置。

【請求項 19】

第四の記憶ボリュームを備え、
通信可能に接続されている第一の記憶装置に、第三の通信規約に従いデータの読み取りを要求する読み取り要求部を備える記憶装置において、

前記第一の命令が、前記第一の記憶装置が備える第三の記憶ボリュームを主記憶ボリューム、前記第四の記憶ボリュームを副記憶ボリュームとするペアを形成する命令である場合に、

前記第三の記憶ボリュームを主記憶ボリューム、前記第四の記憶ボリュームを副記憶ボリュームとする前記ペアが形成されていることをペア管理情報として記憶するペア管理部と、

前記第三の記憶ボリュームのデータを前記第三の通信規約に従い前記第四の記憶ボリュームに読み取る要求を、前記読み取り要求部に送信する形成コピー部と
を備えることを特徴とする請求項 18 に記載の記憶装置。

【請求項 20】

第四の記憶ボリュームと第六の記憶ボリュームとを備える記憶装置において、
前記第一の命令が、前記第一の記憶装置が備える第三の記憶ボリュームを副記憶ボリューム、前記第四の記憶ボリュームを主記憶ボリュームとするペアを形成する命令である場合に、

前記第三の記憶ボリュームを副記憶ボリューム、前記第四の記憶ボリュームを主記憶ボリュームとする前記ペアが形成されていることをペア管理情報として記憶するペア管理部と、

前記第四の記憶ボリュームに書き込まれるデータの複製及び位置情報を前記第六の記憶ボリュームにジャーナルとして記憶するジャーナル記憶部と
を備えることを特徴とする請求項 18 に記載の記憶装置。

【請求項 21】

第六の記憶ボリュームを備え、
前記第三の記憶ボリュームを副記憶ボリューム、前記第四の記憶ボリュームを主記憶ボリュームとするペアが形成されている記憶装置において、

前記第一の命令が、前記第三の記憶ボリュームを主記憶ボリューム、前記第四の記憶ボリュームを副記憶ボリュームとする前記ペアの交換命令である場合に、

前記ペア管理情報に、前記第三の記憶ボリュームを主記憶ボリューム、前記第四の記憶ボリュームとして記憶する前記ペア管理部と、

前記第四の記憶ボリュームに書き込まれるデータの履歴を前記第六の記憶ボリュームにジャーナルとして記憶する処理を停止するジャーナル停止部と
を備えることを特徴とする請求項 20 に記載の記憶装置。

【書類名】明細書**【発明の名称】** 記憶装置システムの制御方法、記憶装置システム、及び記憶装置**【技術分野】****【0001】**

本発明は、記憶装置システムの制御方法、記憶装置システム、及び記憶装置に関する。

【背景技術】**【0002】**

情報処理システムにおける災害復旧（ディザスタリカバリ）が注目されている。ディザスタリカバリを実現する技術として、プライマリサイトに設置されている記憶装置のデータの複製を、これとは遠隔したリモートサイトに設置されている記憶装置においても管理する技術が知られている。プライマリサイトの被災時に、リモートサイトに設置されている記憶装置のデータを使用することで、プライマリサイトで行われていた処理をリモートサイトで継続して行うことができる。

【0003】

プライマリサイトからリモートサイトへのデータ転送の方法としては、プライマリサイトの情報処理装置とリモートサイトの情報処理装置との間でデータを授受する方法が知られている。プライマリサイトの情報処理装置は、プライマリサイトの記憶装置に書き込まれたデータの複製を、リモートサイトの情報処理装置に送信する。これを受信したリモートサイトの情報処理装置がリモートサイトの記憶装置に当該データの書き込み要求を送信する。

【発明の開示】**【発明が解決しようとする課題】****【0004】**

前述の方法によるデータのバックアップでは、情報処理装置間のネットワーク上を大量のデータが流れる。そのため、情報処理装置のインタフェース処理負荷の増大や、情報処理装置間で行われる他のデータ伝送の遅延等が問題となっていた。また、この方法では、各情報処理装置にデータのバックアップを制御するためのソフトウェアをインストールする必要がある。そのため、データのバックアップを行っている全ての情報処理装置に対して、当該ソフトウェアのバージョンアップ等の管理作業を行う必要があり、管理コストが増大している。

【0005】

本発明は上記課題を鑑みてなされたものであり、記憶装置システムの制御方法、記憶装置システム、及び記憶装置を提供することを目的とする。

【課題を解決するための手段】**【0006】**

上記課題を解決するための本発明の主たる発明は、
情報処理装置と、
第一の記憶ボリュームを備える第一の記憶装置と、
第二の記憶ボリュームを備える第二の記憶装置と
を備え、
前記情報処理装置と前記第一の記憶装置とが通信可能に接続され、
前記第一の記憶装置と前記第二の記憶装置とが通信可能に接続され、
前記情報処理装置が、第一の通信規約に従い前記第一の記憶装置にデータの書き込みを要求する第一の書き込み要求部を備え、
前記第一の記憶装置が、第二の通信規約に従い第二の記憶装置にデータの書き込みを要求する第二の書き込み要求部を備える
記憶装置システムの制御方法において、
前記情報処理装置が、前記第二の記憶装置で実行される第一の命令を第一のデータに設定するステップと、
前記情報処理装置が、前記第一のデータを前記第一の記憶ボリュームに書き込む要求を

、前記第一の書き込み要求部に送信するステップと、

前記第一の記憶装置が、前記第一の記憶ボリュームに書き込まれている前記第一のデータが前記第二の記憶装置に対する命令である場合に、前記第一のデータを前記第二の記憶ボリュームに書き込む要求を、前記第二の書き込み要求部に送信するステップと、

前記第二の記憶装置が、前記第二の記憶ボリュームに書き込まれている前記第一のデータに設定されている前記第一の命令を実行するステップとを備えることとする。

【0007】

ここで、情報処理装置とは、例えば、パーソナルコンピュータ、ワークステーション、メインフレームコンピュータなどである。記憶装置とは、例えば、ディスクアレイ装置や半導体記憶装置などである。記憶ボリュームとは、ディスクドライブにより提供される物理的な記憶領域である物理ボリュームと、物理ボリューム上に論理的に設定される記憶領域である論理ボリュームとを含む記憶資源である。また、通信規約とは、例えば SCS I (Small Computer System Interface) 規格により定められる WR I T E コマンドである。そのため、情報処理装置がオペレーティングシステムに新たな命令を追加することなく、第二の記憶装置に第一の命令を実行させることができるようになる。

【0008】

ここで、例えば、第一の命令を、第一の記憶装置のデータの読み取り命令とすれば、情報処理装置からの指示により、第二の記憶装置に第一の記憶装置のデータの複製を持つことができる。この方法により、データのバックアップ運用において、情報処理装置間でのデータ通信量を減らすことができる。また、データのバックアップを制御するためのソフトウェアを、データのバックアップを行っている全ての情報処理装置にインストールする必要も無くなり、管理コストを減らすことができる。

【0009】

その他、本願が開示する課題、及びその解決方法は、発明を実施するための最良の形態の欄、及び図面により明らかにされる。

【発明の効果】

【0010】

記憶装置システムの制御方法、記憶装置システム、及び記憶装置を提供することができる。

【発明を実施するための最良の形態】

【0011】

===全体構成例===

まず、本実施の形態に係る記憶装置システムを含む情報処理システムの概略構成を図 1 に示す。本実施の形態に係る情報処理システムは、情報処理装置 11、第一の記憶装置 10、第二の記憶装置 20 を含んで構成されている。また、第一の記憶装置 10 は、第一の記憶装置がデータ入出力処理を行う論理ボリューム 30 (以後、「第一の論理ボリューム」と称する) を備え、第二の記憶装置 20 は、第二の記憶装置がデータ入出力処理を行う論理ボリューム 40 (以後、「第二の論理ボリューム」と称する) を備えている。

【0012】

また、情報処理装置 11 と第一の記憶装置 10 とは、第一のネットワーク 50 により通信可能に接続されている。第一のネットワーク 50 は、例えば、L A N (Local Area Network) や S A N (Storage Area Network)、i S C S I (Internet Small Computer System Interface)、E S C O N (Enterprise Systems Connection) (登録商標)、F I C O N (Fibre Connection) (登録商標) などである。

【0013】

第一の記憶装置 10 と第二の記憶装置 20 とは、第二のネットワーク 60 により通信可能に接続されている。第二のネットワーク 60 は、例えば、ギガビットイーサネット (登録商標)、A T M (Asynchronous Transfer Mode)、公衆回線などである。

【0014】

=== 情報処理装置 ===

情報処理装置 11 は、CPU (Central Processing Unit) やメモリを備えるコンピュータであり、パーソナルコンピュータやワークステーション、メインフレームなどのコンピュータである。情報処理装置 11 は、結合された複数台のコンピュータで構成されることもある。情報処理装置 11 ではオペレーティングシステムが動作している。オペレーティングシステム上ではアプリケーションソフトウェアが動作している。

【0015】

=== 記憶装置 ===

図 2 に第一及び第二の記憶装置 10, 20 の一例として説明するディスクアレイ装置の具体的な構成を示している。なお、第一及び第二の記憶装置 10, 20 は、ディスクアレイ装置以外にも、例えば、半導体記憶装置などであってもよい。ディスクアレイ装置は、チャンネル制御部 201、リモート通信インタフェース 202、ディスク制御部 203、共有メモリ 204、キャッシュメモリ 205、これらの間を通信可能に接続するクロスバススイッチなどで構成されるスイッチング制御部 206、管理端末 207、及び記憶デバイス 208などを備えて構成される。

【0016】

キャッシュメモリ 205 は、主としてチャンネル制御部 201 とディスク制御部 203 との間で授受されるデータを一時的に記憶するために用いられる。例えばチャンネル制御部 201 が情報処理装置 11, 21 から受信したデータ入出力コマンドが書き込みコマンドである場合には、チャンネル制御部 201 は情報処理装置 11, 21 から受信した書き込みデータをキャッシュメモリ 205 に書き込む。またディスク制御部 203 はキャッシュメモリ 205 から書き込みデータを読み出して記憶デバイス 208 に書き込む。

【0017】

ディスク制御部 203 は、チャンネル制御部 201 により共有メモリ 204 に書き込まれたデータ I/O 要求を読み出してそのデータ I/O 要求に設定されているコマンド (例えば、SCSI 規格のコマンド) に従って記憶デバイス 208 にデータの書き込みや読み出しなどの処理を実行する。ディスク制御部 203 は記憶デバイス 208 から読み出したデータをキャッシュメモリ 205 に書き込む。またデータの書き込み完了通知や読み出し完了通知などをチャンネル制御部 201 に送信する。ディスク制御部 203 は、記憶デバイス 208 をいわゆる RAID (Redundant Array of Inexpensive Disks) 方式に規定される RAID レベル (例えば、0, 1, 5) で制御する機能を備えることもある。

【0018】

記憶デバイス 208 は、例えば、ハードディスク装置である。記憶デバイス 208 はディスクアレイ装置と一体型とすることもできるし、別体とすることもできる。各サイトの記憶デバイス 208 により提供される記憶領域は、この記憶領域上に論理的に設定されるボリュームである論理ボリューム 209 を単位として管理されている。記憶デバイス 208 へのデータの書き込みや読み出しは、論理ボリューム 209 に付与される識別子である LUN (Logical Unit Number) を指定して行なうことができる。また、論理ボリューム 209 は、例えば 512 KB といった決まった単位で管理され、この単位でデータの出入力が行われる。ここで、当該単位を論理ブロックと呼び、各論理ブロックには、位置情報を示す論理ブロックアドレス (以後、「LBA」と称する) が付与されている。

【0019】

管理端末 207 はディスクアレイ装置や記憶デバイス 208 を保守・管理するためのコンピュータである。チャンネル制御部 201 やディスク制御部 203 において実行されるソフトウェアやパラメータの変更は、管理端末 207 からの指示により行われる。管理端末 207 はディスクアレイ装置に内蔵される形態とすることもできるし、別体とすることもできる。

【0020】

リモート通信インタフェース 202 は、他の記憶装置 10, 20 とデータ伝送をするための通信インタフェース (チャンネルエクステンダ) であり、後述するリモートコピーにお

ける複製データの伝送はこのリモート通信インタフェース 202 を介して行われる。リモート通信インタフェース 202 は、チャンネル制御部 201 のインタフェース（例えば、ESCON（登録商標）、FICON（登録商標）などのインタフェース）を第一のネットワーク 40 の通信方式に変換する。これにより他の記憶装置 10, 20 との間でのデータ伝送が実現される。

【0021】

なお、ディスクアレイ装置は、以上に説明した構成のもの以外にも、例えば、NFS（Network File System）などのプロトコルにより情報処理装置 11 からファイル名指定によるデータ入出力要求を受け付けるように構成されたNAS（Network Attached Storage）として機能するものなどであってもよい。

【0022】

共有メモリ 204 はチャンネル制御部 201 とディスク制御部 203 の両方からアクセスが可能である。データ入出力要求コマンドの受け渡しに利用される他、記憶装置 10, 20, 30 や記憶デバイス 208 の管理情報等が記憶される。本実施の形態においては、図 3 に示すLUNマップ情報テーブル 301、図 5 に示すコマンドデバイス管理テーブル 501、図 10 に示すペア管理テーブル 1001 が共有メモリ 204 に記憶される。

【0023】

===仮想ボリューム===

前述の通り、論理ボリューム 209 は物理ボリューム上に論理的に設定される記憶領域である。また、論理ボリュームとして「仮想ボリューム」を用いることにより、論理ボリューム 209 が設定されている記憶装置 10, 20 と、当該論理ボリューム 209 と対応づけられた物理ボリュームが備えられた記憶装置 10, 20 とを別のものとすることができる。

【0024】

この機能を実現するため、第一の記憶装置 10 は、図 3 に示すLUNマップ情報テーブル 301 を記憶している。このLUNマップ情報テーブル 301 には、第一の記憶装置 10 が取り扱う論理ボリューム 209 に関する情報が記述されている。

【0025】

LUNの欄には、各論理ボリュームのLUNが記述される。論理ボリューム 209 が仮想ボリュームである場合は、ターゲットの欄に、当該仮想ボリュームと対応づけられている論理ボリューム 209 が備えられている記憶装置が設定される。さらに、マッピングLUNの欄には、仮想ボリュームと対応づけられている論理ボリューム 209 のLUNが設定される。つまり、マッピングLUNの欄に記述のある場合は、その論理ボリュームが仮想ボリュームであることを意味している。

【0026】

LUNマップ情報テーブル 301 の内容は、例えば、第一の記憶装置 10 に接続された管理端末 207 などからオペレータにより登録される。

【0027】

第一の記憶装置 10 は、以上に説明したLUNマップ情報テーブル 301 を用い、後述する仕組みにより、第二の記憶装置 20 の第二の論理ボリューム 40 を、あたかも第一の記憶装置 10 の第一の論理ボリューム 30 であるかのように、情報処理装置 11 に提供する。つまり、情報処理装置 11 は、第二の記憶装置 20 の論理ボリューム 209 に対するデータ入出力要求を、第一の記憶装置 10 に対して行うことができる。

【0028】

情報処理装置 11 から送信されたデータ入出力要求が、データ書き込み要求である場合における、記憶装置システムの処理を図 4 を用いて説明する。情報処理装置 11 は、第一の通信規約に従い第一の記憶装置にデータの書き込みを要求する第一の書き込み要求部 401 を備えている。第一の記憶装置 10 は、第一の書き込み要求部 401 からデータ書き込み要求を受信すると（S401）、このデータ書き込み要求とともに受信した書き込み対象データを、キャッシュメモリ 205 に記憶する。

【0029】

第一の記憶装置10が備えるデータ転送部402は、LUNマップ情報テーブル301を参照し、書き込み要求に設定された第一の論理ボリューム30に対して、マッピングLUNが設定されているかどうかを確認する。マッピングLUNに第二の論理ボリューム40が設定されていると、データ転送部402は、当該データを第二の通信規約に従い第二の論理ボリューム40に書き込む要求を第二の書き込み要求部403に送信する。ここで、第二の書き込み要求部403とは、第二の通信規約に従い第二の記憶装置20にデータの書き込みを要求するものである。第二の記憶装置20は、第二の書き込み要求部403から、当該データの書き込み要求を受信し、当該データを第二の論理ボリューム40に書き込む(S402)。

【0030】

なお、第一の通信規約、及び第二の通信規約とは、例えばSCSI規約により定められるWRITEコマンドである。従って、第一の記憶装置10、及び第二の記憶装置20におけるデータ書き込みインタフェースを変更する必要が無い。

【0031】

以上は書き込み処理についての説明であったが、論理ボリュームからのデータの読み出し処理についても、以上の書き込みの場合とデータの転送の向きが逆になるだけで、それ以外は同様の仕組みで行われる。

【0032】

以上に説明したように、この記憶装置システムにおいては、情報処理装置11は、それがあたかも第一の記憶装置10上の論理ボリュームであるかのように第二の論理ボリュームにアクセスする。

【0033】

===コマンドデバイス===

記憶装置10、20は、特殊な命令を制御するための「コマンドデバイス」を備えている。コマンドデバイスは、情報処理装置11から記憶装置10、20への命令伝達に用いられ、記憶装置10、20はコマンドデバイスに記憶されている命令を実行することができる。通常の命令と異なるのは、コマンドデバイスが論理ボリューム209であることである。このコマンドデバイスの機能について、説明を行う。

【0034】

図5は、記憶装置10、20が記憶しているコマンドデバイス管理テーブル501である。装置の欄は、コマンドデバイスがどの記憶装置10、20に対するものであるかを示している。コマンドデバイスLUNの欄には、コマンドデバイスの実体を表す論理ボリューム209のLUNが設定されている。コマンドデバイス管理テーブル501の内容は、例えば、記憶装置10、20に接続された管理端末207などからオペレータにより登録される。

【0035】

また、各記憶装置10、20のコマンドデバイス管理テーブル601には、他の記憶装置10、20のコマンドデバイスも登録することができる。他の記憶装置10、20のコマンドデバイスを登録する場合、コマンドデバイスLUNには、他の記憶装置10、20のコマンドデバイスLUNに対応する、仮想ボリュームのLUNを登録する。

【0036】

図6は、コマンドデバイスに書き込まれるデータの形式である、コマンドデバイスインタフェース601を表している。コマンドデバイスインタフェース601は、制御パラメタ、入力パラメタ、及び編集データから構成されている。制御パラメタは、記憶装置が実行する命令を示す「処理番号」と、当該命令を実行した結果、データの出力があるかどうかを示す「編集データ有無」で構成されている。入力パラメタには、当該命令を実行する際に用いられるパラメタ情報が設定される。また、編集データには、当該命令を実行した結果出力されたデータが設定される。

【0037】

コマンドデバイスを用いて命令を実行する流れの概要を、図7を用いて説明する。情報処理装置11は、命令設定部701と命令送信部702とを備える。命令設定部701は、第一の記憶装置10で実行される命令の「処理番号」と「編集データ有無」をコマンドインタフェース601に設定したデータを生成する。命令送信部702は、当該データを第一の通信規約に従い第一の記憶装置10のコマンドデバイスである第一の論理ボリュームに書き込む要求を第一の書き込み要求部401に送信する。

【0038】

第一の記憶装置10は、命令実行部703を備える。また、命令実行部703は、後述する論理ボリューム209のペアを制御する、ペア管理部704、形成コピー部705、リストア部706、ジャーナル記憶部707、ジャーナル取得部708、ジャーナル停止部709を備える。

【0039】

命令実行部703は、コマンドデバイス管理テーブル501を参照し、第一の記憶装置10に対するコマンドデバイスLUNを取得する(S701)。命令実行部703は、当該コマンドデバイスを参照し(S702)、コマンドデバイスインタフェース601形式のデータが存在すれば、当該データの処理番号で指定された命令を実行する。

【0040】

図8及び図9のフローチャートを用いて、情報処理装置11、及び記憶装置10、20の処理の流れを説明する。まず、情報処理装置11はコマンドデバイスインタフェース601の形式の第一のデータに、処理番号と編集データ有無とを設定する(S801)。次に、情報処理装置11は、記憶装置10、20に記憶されているコマンドデバイス管理テーブル501を参照し、命令を実行する記憶装置10、20のコマンドデバイスLUNを取得する。情報処理装置11は、記憶装置10、20に対して、第一のデータを当該コマンドデバイスに書き込むWrite要求を送信する(S802)。記憶装置10、20は当該Write要求を受信すると、第一のデータをコマンドデバイスに書き込む。なお、コマンドデバイスは論理ボリューム209であり、当該Write要求は、コマンドデバイス以外の論理ボリューム209に対する書き込み要求と同じ通信規約に基づいて送信される。

【0041】

記憶装置10、20は、コマンドデバイス管理テーブル501を参照し、自身の記憶装置10、20のコマンドデバイスにデータがあるかどうか監視している(S901)。コマンドデバイスに第一のデータが存在していると、記憶装置10、20は、第一のデータの処理番号で指定された命令を実行する(S902)。記憶装置10、20は、当該命令の実行が完了すると第一のデータの編集データ有無を確認する(S903)。記憶装置10、20は、編集データが無い場合は、第一のデータをコマンドデバイスから削除する(S906)。記憶装置10、20は、編集データがある場合は、当該命令を実行し出力されたデータを編集データに設定する(S904)。

【0042】

情報処理装置11は、当該命令の編集データ有無を確認し(S803)、編集データがある場合は、第一のデータの編集データを読み取るRead要求を、記憶装置10、20に対して送信する(S804)。情報処理装置11は、記憶装置10、20から編集データを受領すると(S805)、処理を終了する。なお、当該Read要求は、コマンドデバイス以外の論理ボリューム209に対する読み出し要求と同じ通信規約に基づいて送信される。

【0043】

編集データがある場合、記憶装置10、20は、情報処理装置11から当該編集データに対するRead要求を受領した後に(S905)、第一のデータをコマンドデバイスから削除する(S906)。

【0044】

このように、情報処理装置11が記憶装置10、20の通常の論理ボリューム209に対するデータの読み書きに使用するRead/Write要求を用いて、情報処理装置11が記憶装置10、20に命令を伝達することができる。また、前述の仮想ボリュームを

用いることにより、情報処理装置 11 が、第一の記憶装置 10 を経由して第二の記憶装置 20 に命令を伝達し、第二の記憶装置 20 に命令を実行させることができる。

【0045】

なお、情報処理装置 11 は、以後に説明する「ペア形成」、「ジャーナル取得」、「ジャーナルの処理状況の取得」、「リストア」、及び「スワップ」を記憶装置 10、20 に要求する際に、仮想ボリュームとコマンドデバイスとを使用する。

【0046】

====ペア形成====

次に、本実施の形態において、第一の記憶装置 10 の論理ボリューム 209 のデータの複製を第二の記憶装置 20 の論理ボリューム 209 に記憶させる方法について説明する。

図 10 にペア管理テーブル 1001 を示す。複製元装置の欄には、複製元の論理ボリューム（以後、「主ボリューム」と称する）が属する記憶装置 10、20 が示されている。複製先装置の欄には、複製先の論理ボリューム（以後、「副ボリューム」と称する）が属する記憶装置 10、20 が示されている。また、主 LUN の欄には主ボリュームの LUN が、副 LUN の欄には副ボリュームの LUN が設定されている。この主ボリュームと副ボリュームとの対応を「ペア」と称する。主ボリューム、及び副ボリュームには、後述するジャーナルを格納するための論理ボリューム 209 が割り当てられている。主ジャーナル LUN の欄には、主ボリューム用に割り当てられたジャーナル（以後、「主ジャーナル」と称する）の論理ボリューム 209 の LUN が設定されている。副ジャーナル LUN の欄には、副ボリューム用に割り当てられたジャーナル（以後、「副ジャーナル」と称する）の論理ボリューム 209 の LUN が設定されている。

【0047】

なお、ジャーナルとして論理ボリューム 209 を割り当てる方法は、ユーザ自身がジャーナルとして使用する論理ボリューム 209 を指定しても良いし、情報処理装置 11 が、未使用の論理ボリューム 209 を任意に選択して使用するようにしても良い。

【0048】

図 11 を用いて、ペアを形成する流れを説明する。ここで、第一の記憶装置 10 は第三の論理ボリュームと第五の論理ボリュームとを備え、第二の記憶装置 20 は第四の論理ボリュームと第六の論理ボリュームとを備えているとする。情報処理装置 11 は、第三の論理ボリュームを主ボリューム、第四の論理ボリュームを副ボリューム、第五の論理ボリュームを主ジャーナル、第六の論理ボリュームを副ジャーナルとするペアを形成する命令を第一の記憶装置 10 と第二の記憶装置 20 とに送信する (S1101, S1102)。第一及び第二の記憶装置 10、20 のペア管理部 704 は、当該ペアの状態を各記憶装置 10、20 のペア管理テーブル 1001 に記憶する。第二の記憶装置 20 の形成コピー部 705 は、主ボリュームの Read 要求を第一の記憶装置 10 に送信し、第一の記憶装置 10 から主ボリュームのデータの複製を受信すると、そのデータを副ボリュームに書き込む (S1103)。これにより、主ボリュームのデータと副ボリュームのデータとを一致させることができる。ペア形成指示により、主ボリュームと副ボリュームとを一致させる処理を、「形成コピー」と呼ぶ。

【0049】

また、第一の記憶装置 10 のジャーナル記憶部 707 は、主ボリュームに書き込まれたデータの複製及び位置情報を主ジャーナルに取得する処理を開始する。主ボリュームと主ジャーナルとの対応を、図 12 を用いて説明する。主ジャーナルは、メタデータ領域 1201 とジャーナルデータ領域 1202 とで構成される。第一の記憶装置 10 のジャーナル記憶部 707 は、主ボリュームに書き込まれたデータの複製（以後、「ジャーナルデータ」と称する）を、ジャーナルデータ領域 1202 に格納する。また、第一の記憶装置 10 のジャーナル記憶部 707 は、主ボリュームにおいてデータ 1203 が更新された時刻、データ 1203 の LBA 1204、対応するジャーナルデータ領域の LBA 1206、及び更新データ長を、メタデータ領域 1201 に格納する。なお、副ジャーナルも主ジャーナルと同じく、メタデータ領域 1201 とジャーナルデータ領域 1202 とで構成されて

いる。

【0050】

また、同様の方法を用いることで、情報処理装置11からの指示により、第二の記憶装置20の論理ボリューム209のデータの複製を第一の記憶装置10の論理ボリューム209に記憶させることも可能である。

【0051】

これにより、情報処理装置間でのデータ通信を行うことなく、かつ、情報処理装置11のオペレーティングシステムに新たな命令を追加することなく、プライマリサイトの記憶装置のデータをリモートサイトの記憶装置にバックアップすることができる。また、本実施の形態においては、リモートサイトの記憶装置がプライマリサイトの記憶装置にRead要求を送信することにより形成コピーを行っている。これにより、形成コピー時におけるプライマリサイトの記憶装置の処理負荷を軽減させている。つまり、プライマリサイトの記憶装置がリモートサイトの記憶装置にデータをWriteする方法においては、プライマリサイトの記憶装置は、リモートサイトの記憶装置がペア形成の準備ができているか確認した後に、リモートサイトの記憶装置にデータをWriteする必要がある。そのため、プライマリサイトの記憶装置の処理負荷が高くなり、本番処理を行っているプライマリサイト全体の性能に影響を与えることとなる。本実施の形態においては、プライマリサイトの記憶装置はリモートサイトの記憶装置からのRead要求に対してデータを送信するだけでなく、プライマリサイトの記憶装置の処理負荷が軽減される。

【0052】

====リストア====

前述の形成コピーが行われた後も、第一の記憶装置10は情報処理装置11からの書き込み要求を受け、主ボリュームのデータを更新している。そのため、主ボリュームのデータと副ボリュームのデータが一致していない状態となる。前述の通り、主ジャーナルには形成コピーが行われた以降のジャーナルデータが記憶されている。そこで、第二の記憶装置20が、主ジャーナルに記憶されているデータを副ジャーナルにコピーし、副ジャーナルに記憶されているデータを副ボリュームに書き込むことにより、主ボリュームで行われたデータの更新を、副ボリュームに対しても行うことができる。

【0053】

ここで、第二の記憶装置20が主ジャーナルに記憶されているデータを副ジャーナルにコピーすることを「ジャーナル取得」、副ジャーナルに記憶されたジャーナルデータを副ボリュームに書き込むことを「リストア」と呼ぶ。

【0054】

図13は、ジャーナル取得の流れを示している。情報処理装置11は、第二の記憶装置20に、ジャーナル取得命令を送信する(S1301)。ジャーナル取得命令を受信した第二の記憶装置20のジャーナル取得部708は、ペア管理テーブル1001を参照し、当該ペアの主ジャーナルLUNを取得する。第二の記憶装置20のジャーナル取得部708は、主ジャーナルのRead要求を、第一の記憶装置10に送信し、第一の記憶装置10から主ジャーナルのデータの複製を受信すると、そのデータを副ジャーナルに書き込む(S1302)。

【0055】

次に、図14を用いてリストアの流れを説明する。情報処理装置11は、第二の記憶装置20に、副ジャーナルから副ボリュームへのリストア命令を送信する(S1401)。リストア命令を受信した第二の記憶装置20のリストア部706は、副ジャーナルに記憶されているジャーナルデータを副ボリュームに書き込む。

【0056】

図15は、本実施の形態における、主ジャーナルと副ジャーナルのジャーナルデータ領域1202を示す図である。主ジャーナルと副ジャーナルのジャーナルデータ領域は、同じ先頭LBAと末尾LBAとで定義されている。主ジャーナルのジャーナルデータ領域1202は、ジャーナルデータが格納されているジャーナル格納済み領域1502、150

3、及び1504と、ジャーナルデータが格納されていないパージ済み領域1501とで構成されている。

【0057】

副ジャーナルのジャーナルデータ領域1202は、既に副ボリュームへのリストアに使用されたジャーナルデータが格納されているリストア済み領域1521、副ボリュームへのリストア対象として指定されたジャーナルデータが格納されているリストア中領域1522、リストア対象として指定されていないジャーナルデータが格納されているリード済み領域1523、及びジャーナル取得命令により主ジャーナルからRead中のジャーナルデータが格納されているリード中領域1524で構成される。

【0058】

記憶装置10、20は、ジャーナルデータが作成された時刻順に、先頭LBAから末尾LBAに向かってジャーナルデータをジャーナルデータ領域1202に記憶する。また、記憶装置10、20は、ジャーナルデータが末尾LBAまで達すると、再度先頭LBAに戻ってジャーナルデータを記憶する。つまり、記憶装置10、20は、ジャーナルデータ領域を先頭LBAと末尾LBAとの間でサイクリックに使用する。

【0059】

主ジャーナルを備える第一の記憶装置10は、ジャーナル格納済み領域1502、1503、及び1504の先頭LBAであるジャーナル・アウトLBA1511と、パージ済み領域1501の先頭LBAであるジャーナル・インLBA1512とを記憶している。ここで、ジャーナル・アウトLBAとジャーナル・インLBAとが等しい場合は、主ジャーナルにはジャーナルデータが格納されていないことになる。

【0060】

副ジャーナルを備える第二の記憶装置20は、リストア済み領域1521の最大LBAであるリストア済みLBA、リストア中領域1522の最大LBAであるリストア予定LBA、リード済み領域1523の最大LBAであるリード済みLBA1533、及びリード中領域1534の最大LBAであるリード予定LBAを記憶している。

【0061】

つまり、リストア済みLBA1531とリストア予定LBA1532とが等しい場合は、情報処理装置11から指示されたリストア処理が完了していることになる。また、リード済みLBA1533とリード予定LBA1534とが等しい場合は、情報処理装置10から指示されたジャーナル取得処理が完了していることになる。

【0062】

情報処理装置11は、第一の記憶装置10、及び第二の記憶装置20に対して、ジャーナルの処理状況の取得要求を送信することができる。各記憶装置10、20は、前述の通り、領域の境界を示すLBAの状態を確認し、これに応答する。

【0063】

また、記憶装置10、20は、ジャーナルデータ領域を前述の通りサイクリックに使用しているため、不要となった領域を解放する必要がある。この解放を「パージ」と呼ぶ。記憶装置10、20は、領域の境界を示すLBAの指すアドレスを変更することによりパージを行うことができる。第一の記憶装置10は、主ジャーナルのジャーナル格納済み領域1502、1503、及び1504のうち、第二の記憶装置20が副ジャーナルに取得済みのジャーナル格納済み領域1502をパージすることができる。この場合、第一の記憶装置10は、ジャーナル・アウトLBA1511を、ジャーナル格納済み領域1503の先頭LBAに変更することにより、ジャーナル格納済み領域1502はパージ済み領域1501となる。第二の記憶装置20は、副ジャーナルのリストア済み領域1521をパージされた領域として扱い、ジャーナル取得命令により取得したジャーナルデータを、リストア済み領域1521に格納する。

【0064】

図16のフローチャートを用いて、ジャーナル取得処理、及びリストア処理の流れを説明する。情報処理装置11は、主ジャーナルの処理状況の取得要求を第一の記憶装置10

に送信する (S1601)。情報処理装置 11 は、主ジャーナルの処理状況が設定されたコマンドデバイスの編集データの Read 要求を第一の記憶装置 10 に送信する (S1602)。情報処理装置 11 は、第一の記憶装置からコマンドデバイスの編集データを受領すると (S1603)、ジャーナル・アウト LBA 1511 から、ジャーナル・イン LBA 1512 の一つ前の LBA までの、ジャーナル取得要求を第二の記憶装置 20 に送信する (S1604)。情報処理装置 11 は、副ジャーナルの処理状況の取得要求を第二の記憶装置に送信する (S1605)。情報処理装置 11 は、副ジャーナルの処理状況が設定されたコマンドデバイスの編集データの Read 要求を第二の記憶装置 20 に送信する (S1606)。情報処理装置 11 は、第二の記憶装置からコマンドデバイスの編集データを受領すると (S1607)、編集データに設定されているリード済み LBA 1533 とリード予定 LBA 1534 とを比較し、ジャーナルの取得が完了しているかどうか確認する (S1608)。ジャーナル取得処理が完了していると、情報処理装置 11 は、リード済み LBA 1533 までのリストア要求を第二の記憶装置 20 に送信する (S1609)。その後、情報処理装置 11 は、リード済み LBA 1533 までのジャーナルデータのパージ要求を、第一の記憶装置 10 に送信する (S1610)。情報処理装置 11 は、ジャーナル取得処理、及びリストア処理を繰り返して行う。

【0065】

これにより、情報処理装置間でデータ通信を行うことなく、かつ、情報処理装置のオペレーティングシステムに新たな命令を追加することなく、プライマリサイトの記憶装置において更新されたデータをリモートサイトの記憶装置に反映させることができる。なお、リモートサイトの記憶装置と通信可能に接続された情報処理装置 11 からの指示により、リモートサイトの記憶装置がプライマリサイトの記憶装置からジャーナルを取得し、データをリストアすることもできる。

【0066】

===スワップ===

第一の記憶装置 10 と通信可能に接続された情報処理装置 11 (以後、「第一の情報処理装置」と称する) からの指示により、第一の記憶装置 10 の主ボリュームと第二の記憶装置 20 の副ボリュームとでペアが形成されているとする。ここで、第一の情報処理装置に障害が発生すると、第二の記憶装置 20 と通信可能に接続された情報処理装置 11 (以後、「第二の情報処理装置」と称する) が、前記ペアの副ボリュームを用いて第一の情報処理装置が行っていた処理を継続して行う。この際に、第二の情報処理装置は、前記ペアの主ボリュームと副ボリュームとの関係の入れ替えを行う。つまり、第二の記憶装置 20 が備える論理ボリューム 209 を主ボリューム、第一の記憶装置 10 が備える論理ボリューム 209 を副ボリュームとしたペアが形成されることになる。このようなペア関係の入れ替えを、「スワップ」と呼ぶ。

【0067】

図 17、18 を用いて、ペアのスワップの流れを説明する。図 17 に示すように、第二の情報処理装置 11 は、第一の記憶装置 10 と第二の記憶装置 20 とに、ペアのスワップ命令を送信する (S1701, S1702)。スワップ命令を受信した第一の記憶装置 10 のジャーナル停止部は、主ボリュームに対するジャーナルの記憶を停止する。また、第一の記憶装置 10 のペア管理部 704 は、ペア管理テーブル 1001 の主ボリュームと副ボリュームとを入れ替える。同様に、スワップ命令を受信した第二の記憶装置 20 のペア管理部 704 は、ペア管理テーブル 1001 の主ボリュームと副ボリュームとを入れ替える。第二の記憶装置 20 のジャーナル記憶部 707 は、主ボリュームとなる第二の記憶装置 20 が備える論理ボリューム 209 のジャーナルの記憶を開始する。

【0068】

図 18 は、当該命令を受信した第一の記憶装置 10 と第二の記憶装置 20 とがスワップ処理を実施し、第二の記憶装置 20 の論理ボリューム 209 を主ボリューム、第一の記憶装置 10 の論理ボリューム 209 を副ボリュームとしたペアが形成された状態を表している。

【0069】

第二の情報処理装置、及び各記憶装置10、20で行うスワップ処理の詳細を、図19～図21のフローチャートを用いて説明する。第二の情報処理装置は、前述したジャーナル取得、及びリストアの処理を実行する(S1901)。第二の情報処理装置は、副ジャーナルの処理状況の取得要求を第二の記憶装置20に送信する(S1902)。第二の情報処理装置は、副ジャーナルの処理状況が設定されたコマンドデバイスの編集データのRead要求を第二の記憶装置10に送信する(S1903)。第二の情報処理装置は、第二の記憶装置20からコマンドデバイスの編集データを受領すると(S1904)、編集データに設定されているリストア済みLBA1531とリストア予定LBA1532とを比較し、リストアが完了しているかどうか確認する(S1905)。リストアが完了していると、第二の情報処理装置は、ペアのスワップ要求を第一の記憶装置10と第二の記憶装置20に送信する(S1906, S1907)。ペアのスワップ要求を受信した第一の記憶装置10は、主ボリュームに対して行っていたジャーナルの取得処理を停止し(S2001)、ペア管理テーブル1001の複製元と複製先との関係を入れ替える(S2002)。また、ペアのスワップ要求を受信した第二の記憶装置20は、ペア管理テーブル1001の複製元と複製先との関係を入れ替え(S2101)、第二の記憶装置20の主ボリュームに対するジャーナルの取得処理を開始する(S2102)。

【0070】

ここで、第一の情報処理装置、及び第一の記憶装置10を備えるプライマリサイトと、第二の情報処理装置、及び第二の記憶装置20を備えるリモートサイトとで構成される情報処理システムがあるとする。第一の情報処理装置に障害が発生すると、第二の情報処理装置は、第二の記憶装置20を用いて、プライマリサイトで行われていた本番処理を継続して行う。第二の情報処理装置は、第一の記憶装置10と第二の記憶装置20とに前述のスワップ命令を指示することにより、第二の記憶装置20を本番で使用し、かつ、第二の記憶装置20のデータを第一の記憶装置10にバックアップすることができる。また、第二の記憶装置20のデータが第一の記憶装置10にバックアップされているため、第一の情報処理装置の障害が回復した際には、本番処理の実行を迅速にプライマリサイトに切り替えることも可能である。

【0071】

また、情報処理装置11から記憶装置10、20へのスワップ指示は、情報処理装置11が備えているRead/Writeコマンドを用いて行われるため、情報処理装置11のオペレーティングシステムに新たな命令を追加する必要がない。

【0072】

以上、本実施の形態について説明したが、上記実施例は本発明の理解を容易にするためのものであり、本発明を限定して解釈するためのものではない。本発明は、その趣旨を逸脱することなく、変更、改良され得ると共に、本発明にはその等価物も含まれる。

【図面の簡単な説明】

【0073】

【図1】 本実施の形態に係る情報処理システムの概略構成を示す図である。

【図2】 本実施の形態に係るディスクアレイ装置の構成を示す図である。

【図3】 本実施の形態に係るLUNマップ情報テーブルを示す図である。

【図4】 本実施の形態に係る仮想ボリュームを用いたデータ書き込みの概略を示す図である。

【図5】 本実施の形態に係るコマンドデバイス管理テーブルを示す図である。

【図6】 本実施の形態に係るコマンドデバイスインタフェースを示す図である。

【図7】 本実施の形態に係るコマンドデバイスに設定された命令を実行する概略を示す図である。

【図8】 本実施の形態に係る情報処理装置におけるコマンドデバイスの制御のフローチャートを示す図である。

【図9】 本実施の形態に係る記憶装置におけるコマンドデバイスの制御のフローチャ

ートを示す図である。

【図10】本実施の形態に係るペア管理テーブルを示す図である。

【図11】本実施の形態に係るペア形成の概略を示す図である。

【図12】本実施の形態に係る主ボリュームと主ジャーナルとの関係を示す図である。

【図13】本実施の形態に係るジャーナル取得の概略を示す図である。

【図14】本実施の形態に係るリストアの概略を示す図である。

【図15】本実施の形態に係る主ジャーナル及び副ジャーナルのジャーナルデータ領域を示す図である。

【図16】本実施の形態に係る情報処理装置におけるジャーナル取得及びリストアのフローチャートを示す図である。

【図17】本実施の形態に係るスワップの概略を示す図である。

【図18】本実施の形態に係るスワップが完了した状態の概略を示す図である。

【図19】本実施の形態に係る第二の情報処理装置におけるスワップのフローチャートを示す図である。

【図20】本実施の形態に係る第一の記憶装置におけるスワップのフローチャートを示す図である。

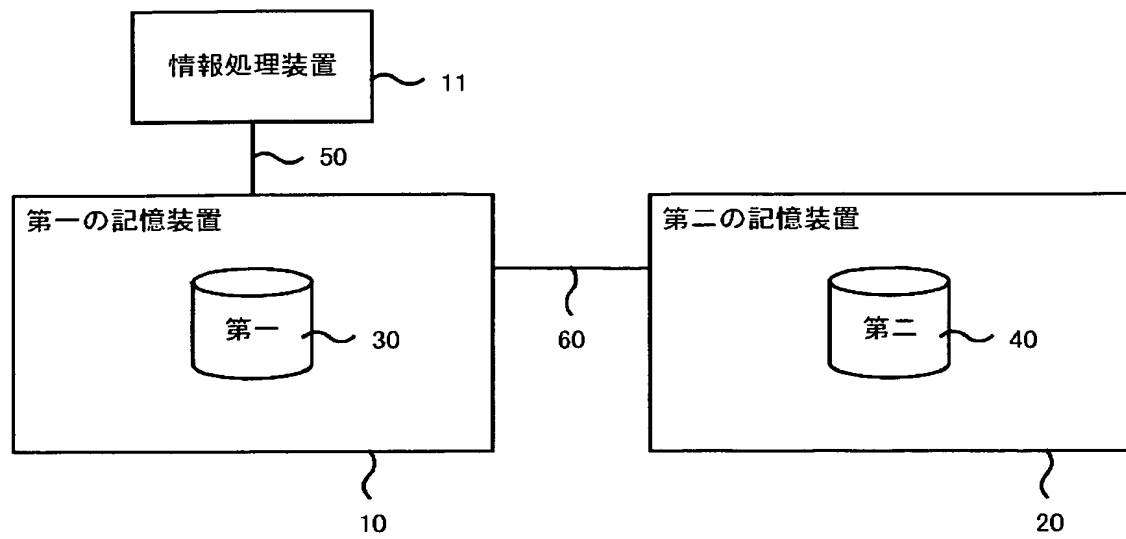
【図21】本実施の形態に係る第二の記憶装置におけるスワップのフローチャートを示す図である。

【符号の説明】

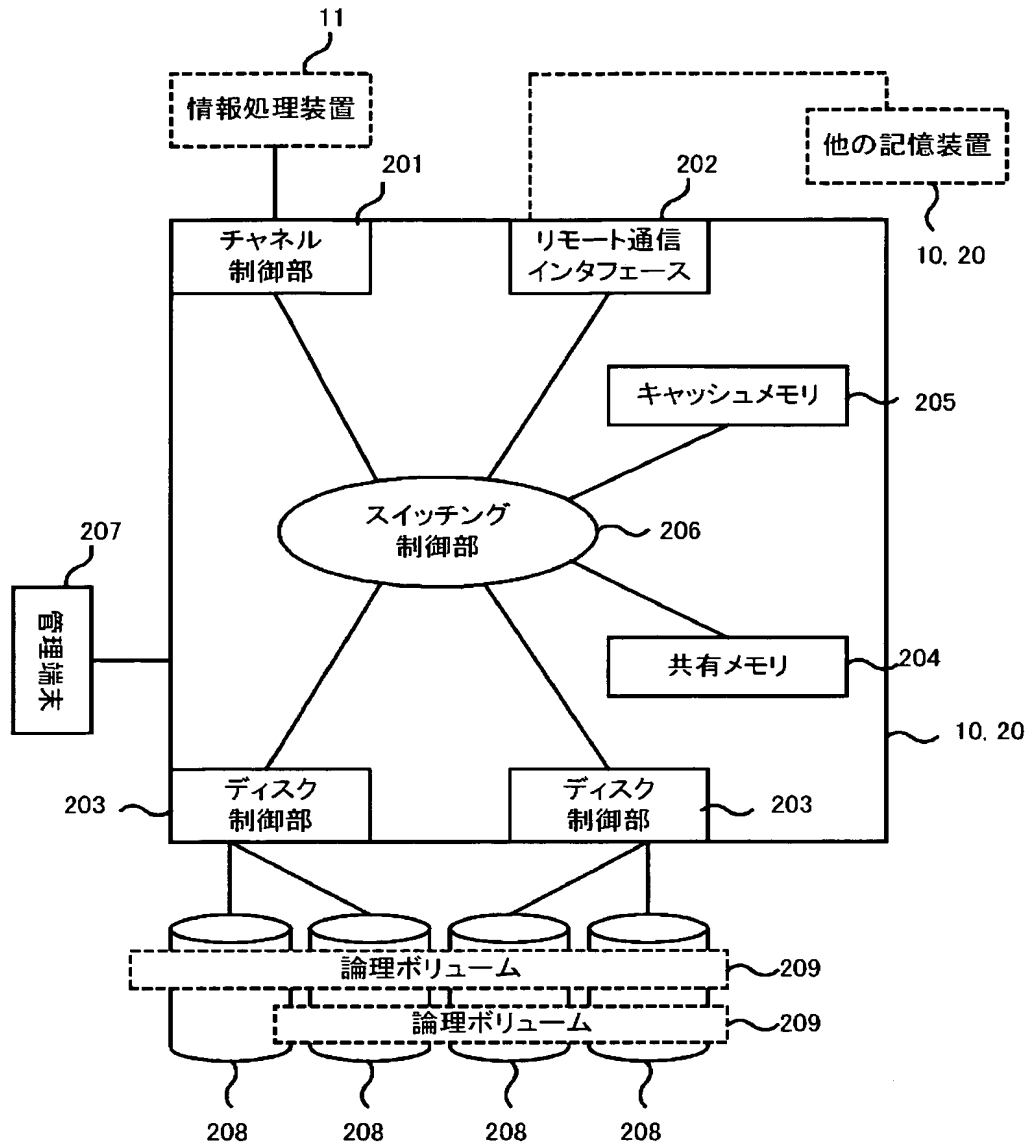
【0074】

10	第一の記憶装置	11	情報処理装置
20	第二の記憶装置	30	第一の論理ボリューム
40	第二の論理ボリューム	201	チャンネル制御部
202	リモート通信インタフェース	203	ディスク制御部
204	共有メモリ	205	キャッシュメモリ
206	スイッチング制御部	207	管理端末
208	記憶デバイス	209	論理ボリューム
301	LUNマップ情報テーブル	401	第一の書き込み要求部
402	データ転送部	403	第二の書き込み要求部
501	コマンドデバイス管理テーブル	601	コマンドデバイスインタフェース
701	命令設定部	702	命令送信部
703	命令実行部	704	ペア管理部
705	形成コピー部	706	リストア部
707	ジャーナル記憶部	708	ジャーナル取得部
709	ジャーナル停止部		
1001	ペア管理テーブル	1101	主ボリューム
1102	副ボリューム	1103	主ジャーナル
1104	副ジャーナル	1201	メタデータ領域
1202	ジャーナルデータ領域	1501	ページ済み領域
1502	ジャーナル格納済み領域	1503	ジャーナル格納済み領域
1504	ジャーナル格納済み領域		
1511	ジャーナル・アウトLBA	1512	ジャーナル・インLBA
1521	リストア済み領域	1522	リストア中領域
1523	リード済み領域	1524	リード中領域
1531	リストア済みLBA	1532	リストア予定LBA
1533	リード済みLBA	1534	リード予定LBA

【書類名】 図面
【図 1】



【図 2】

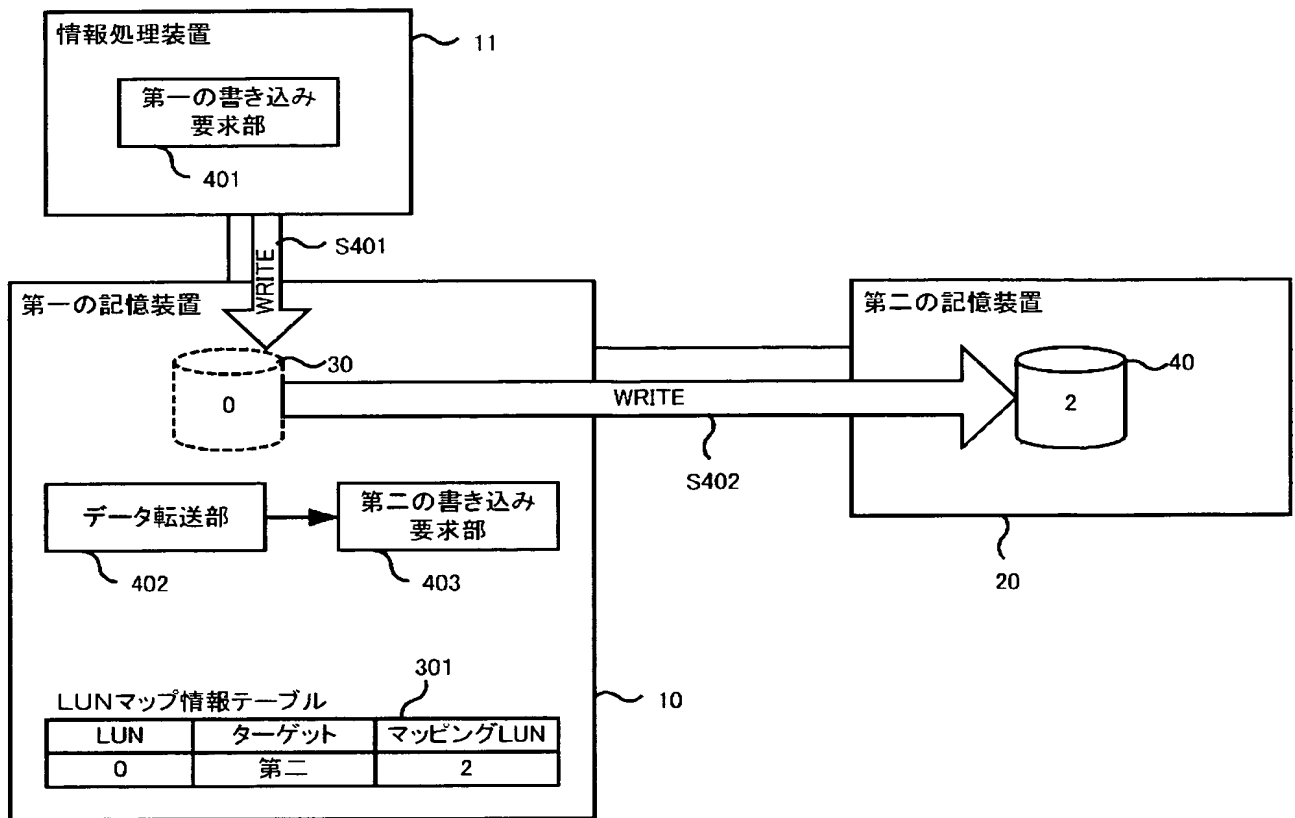


【図 3】

LUN	ターゲット	マッピングLUN
0	第二	0
1	第三	1
⋮	⋮	⋮

301

【図 4】



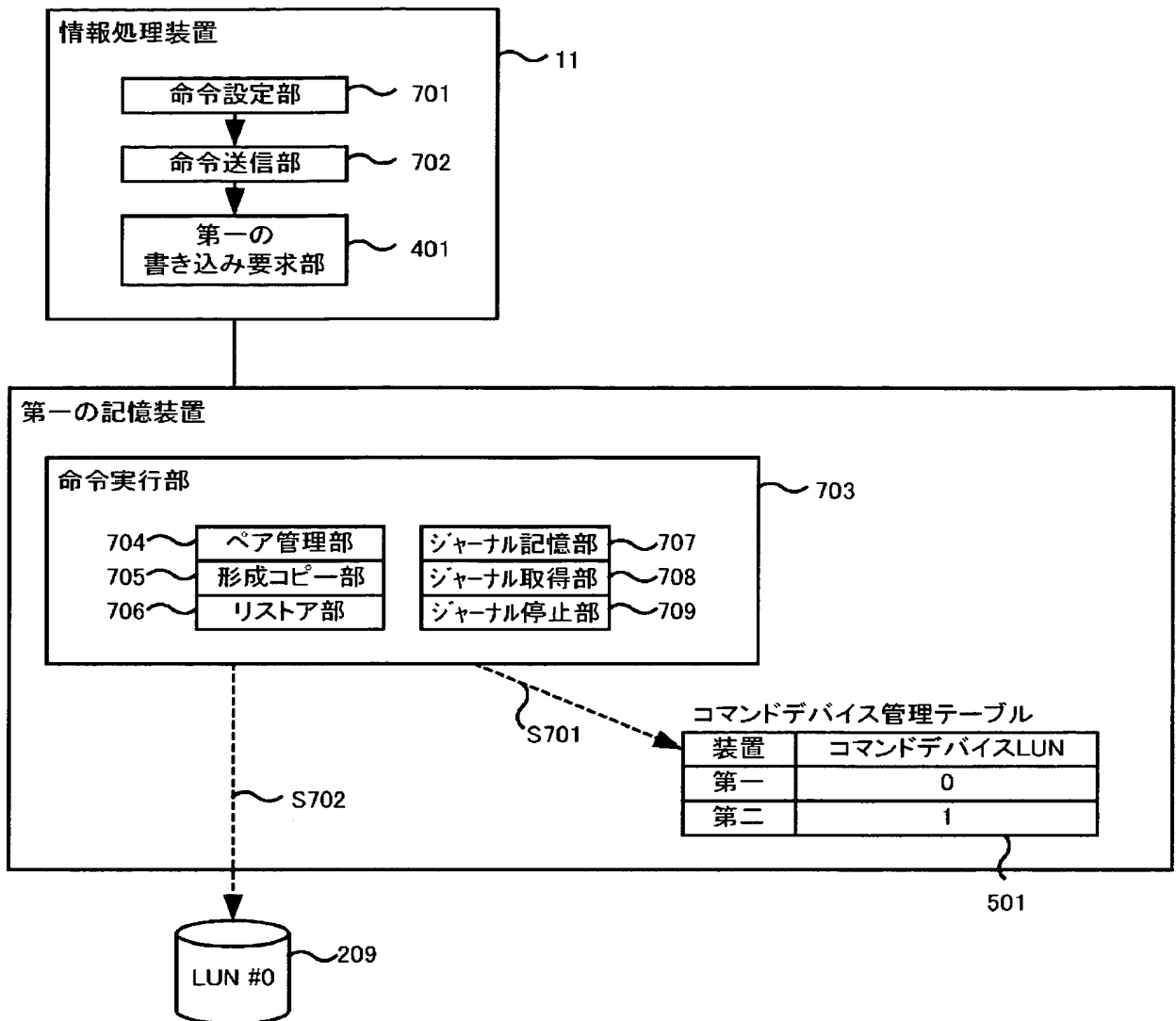
【図 5】

装置	コマンドデバイスLUN
第一	0
第二	1

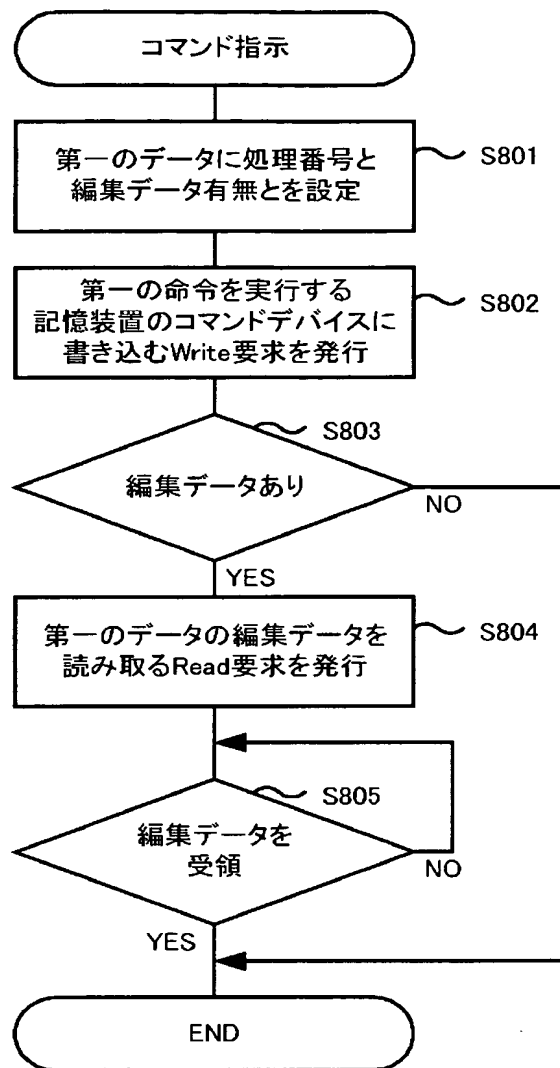
【図 6】

制御パラメタ
処理番号
編集データ有無
入力パラメタ
編集データ

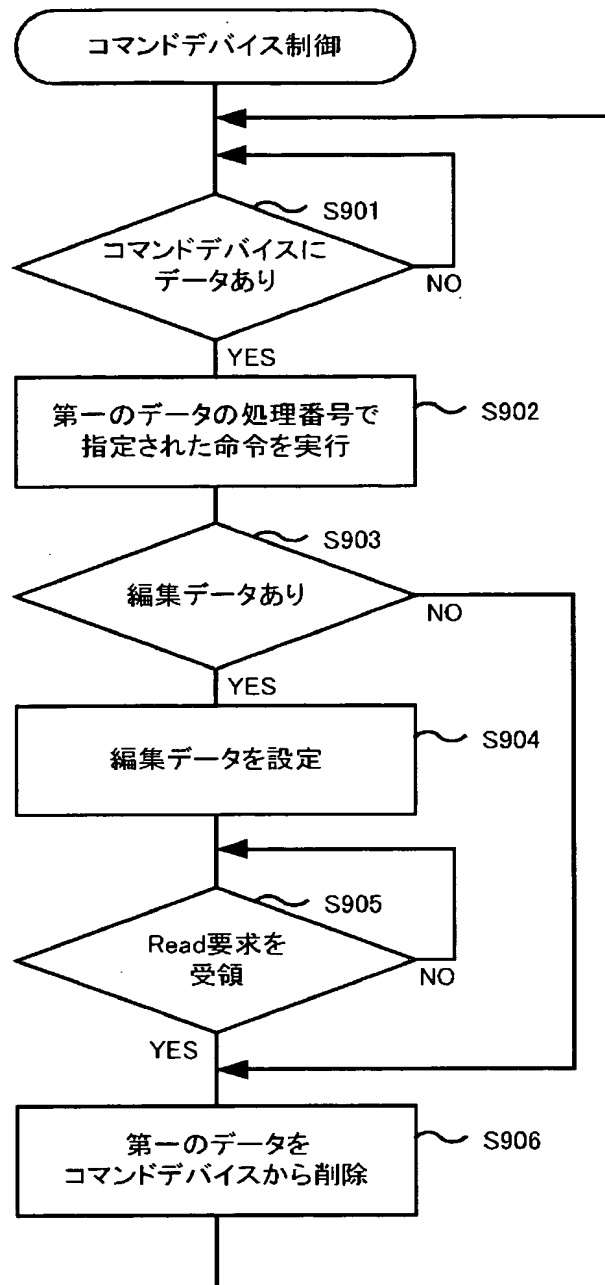
【図 7】



【図 8】



【図 9】

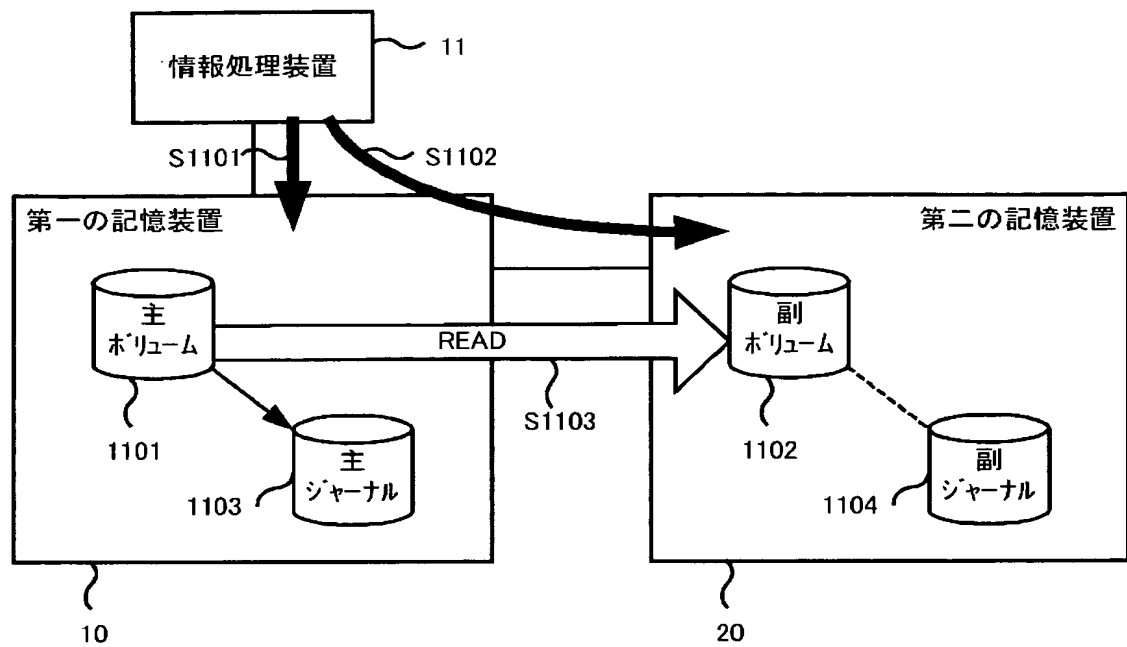


【図 10】

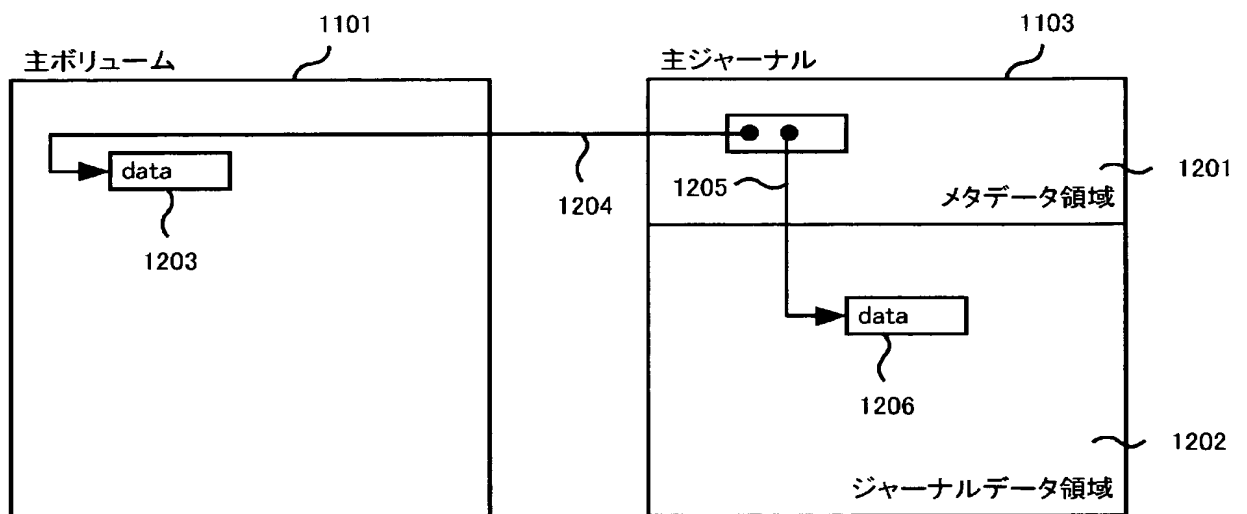
複製元装置	複製先装置	主LUN	主ジャーナルLUN	副LUN	副ジャーナルLUN
第一	第二	0	3	0	3
第一	第二	1	4	1	4
第一	第二	2	5	2	5
⋮	⋮	⋮	⋮	⋮	⋮

1001

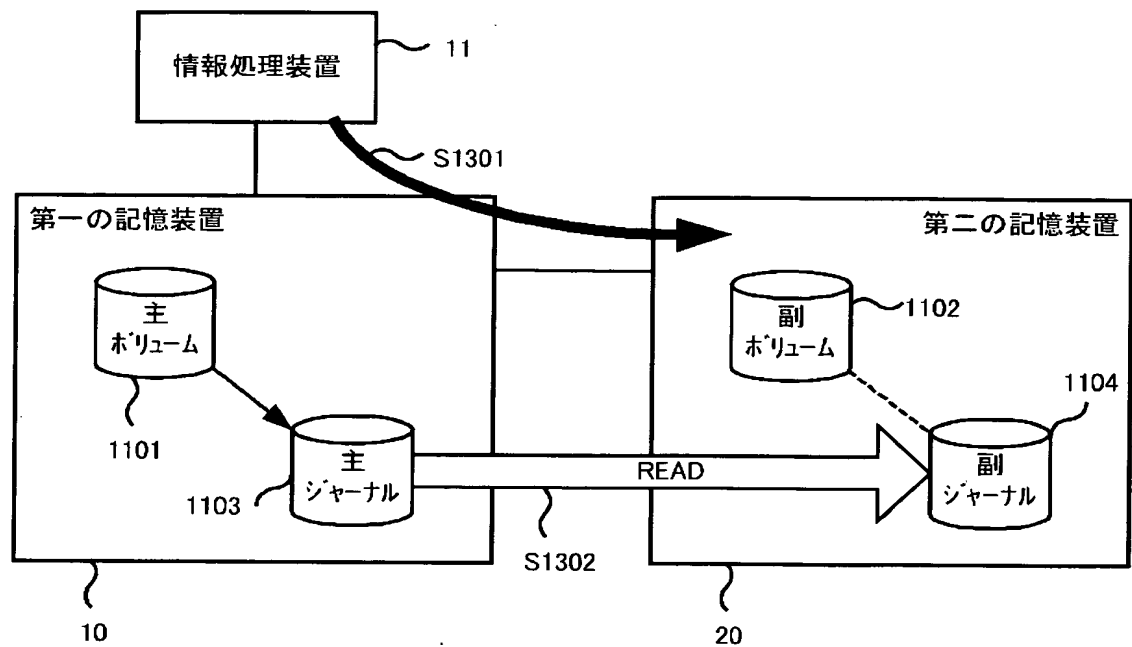
【図 1 1】



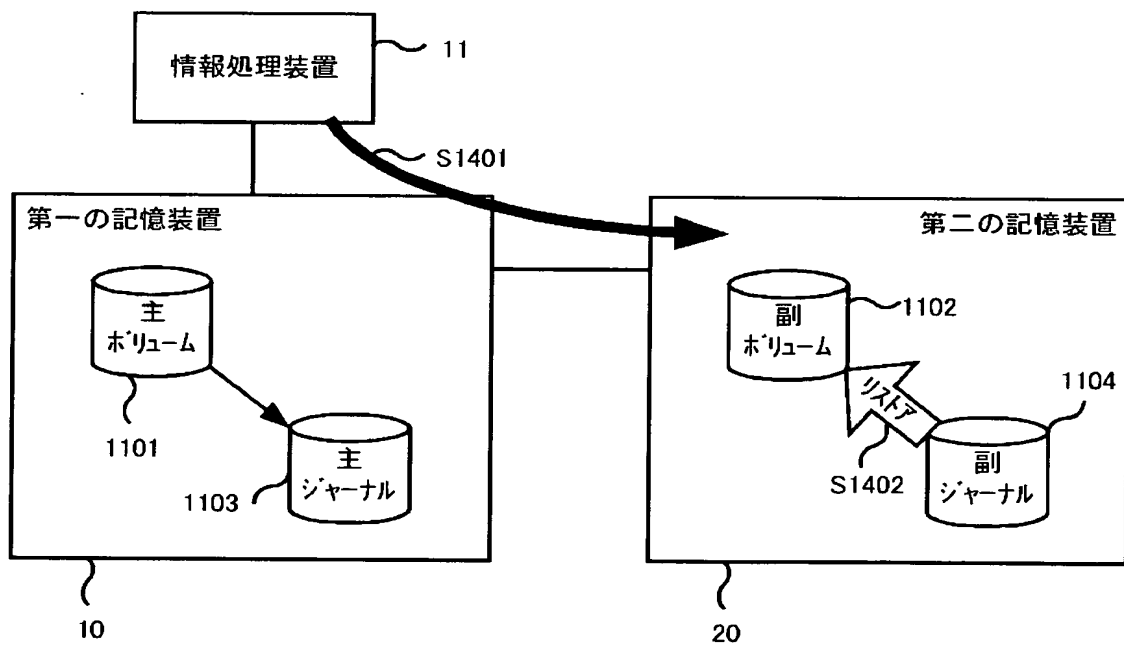
【図 1 2】



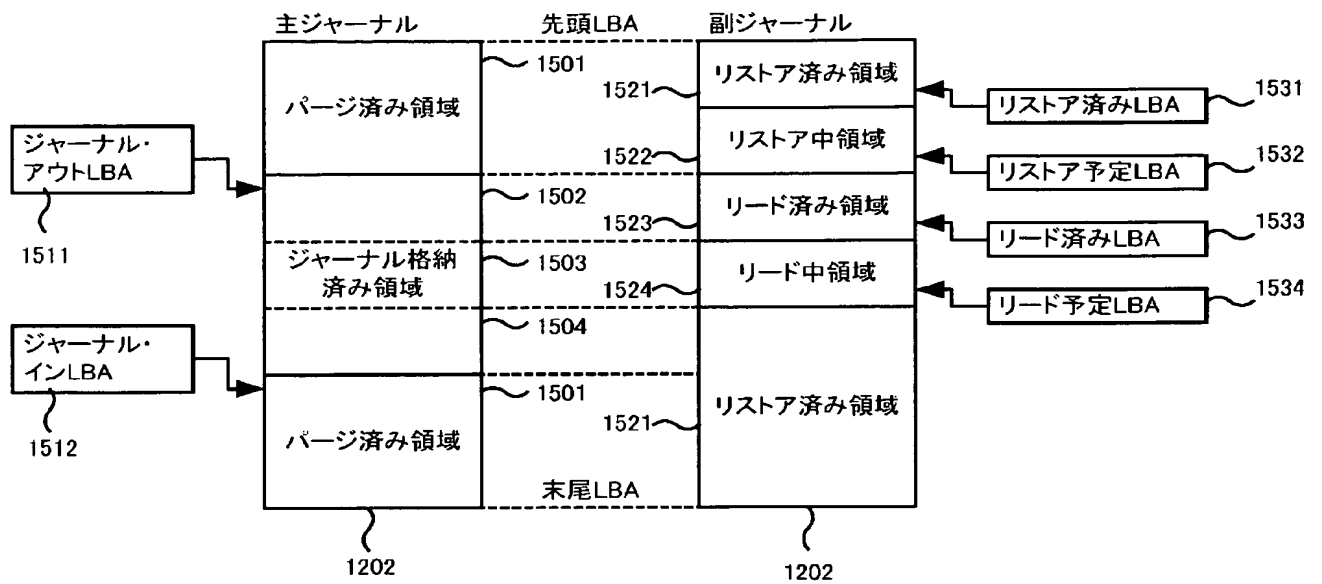
【図 13】



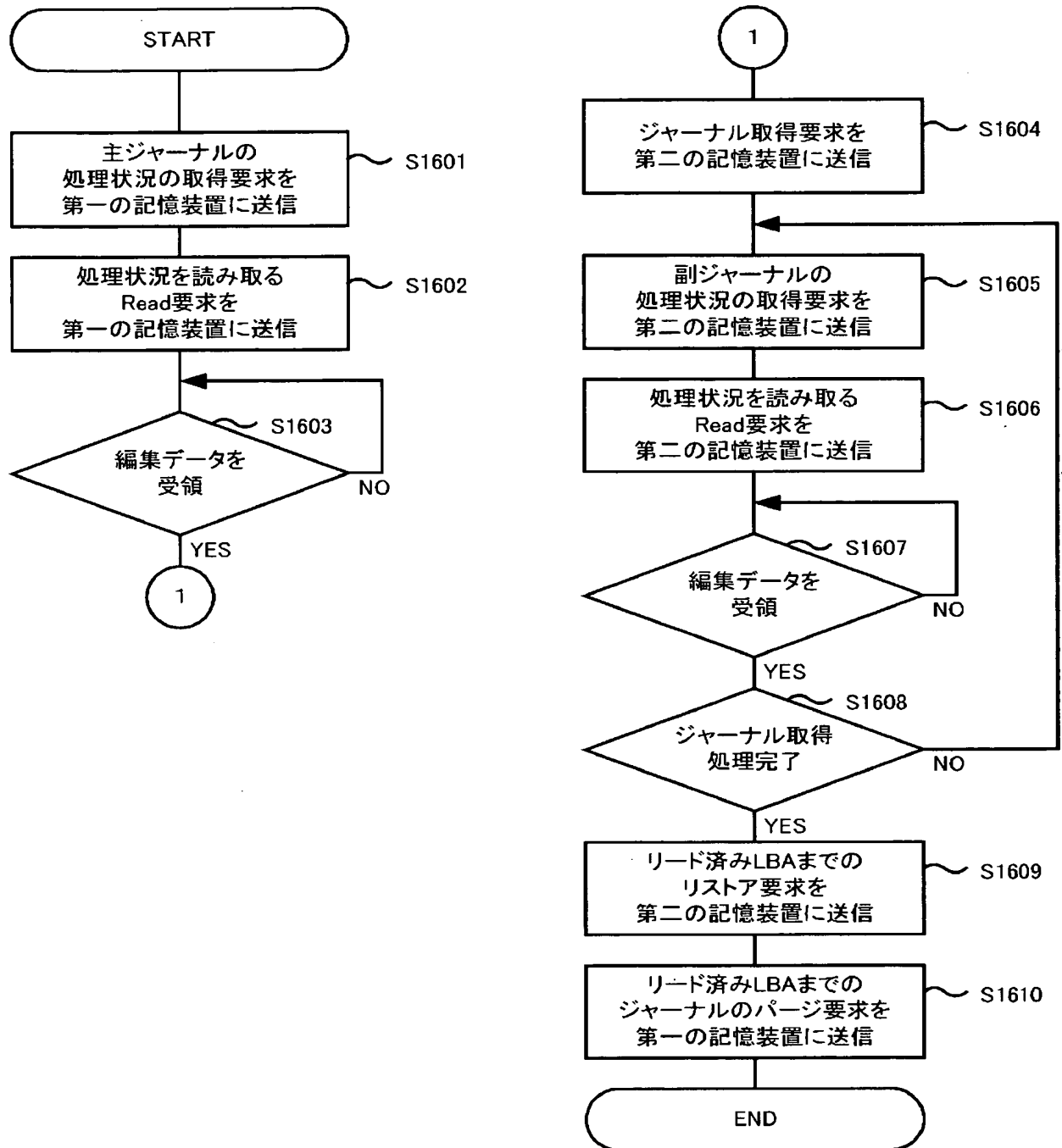
【図 14】



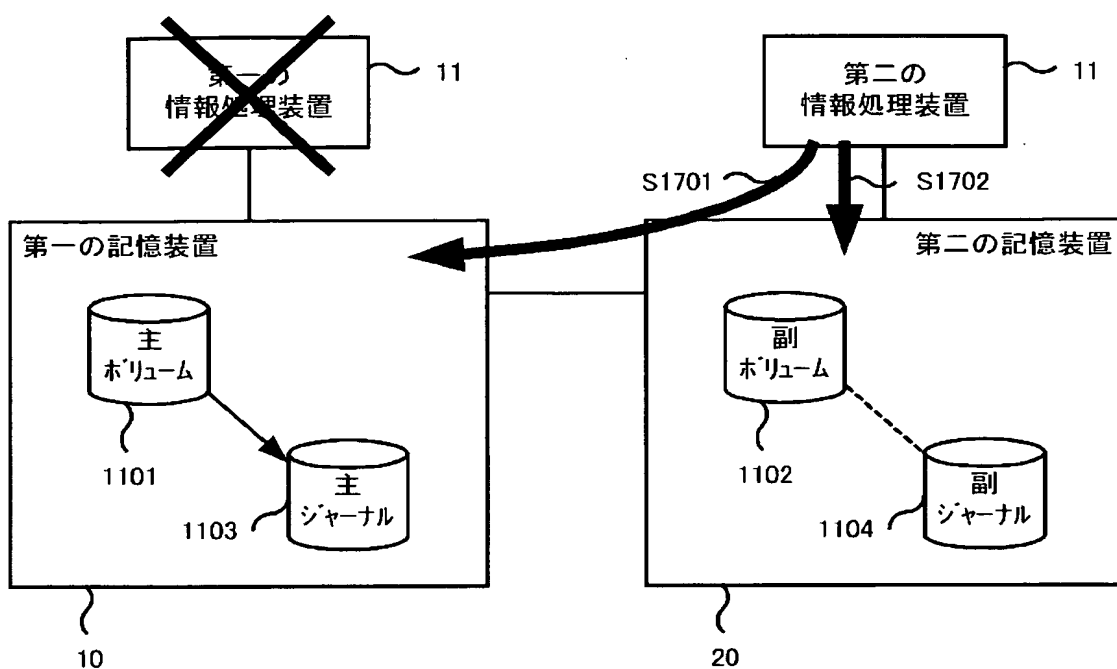
【図 15】



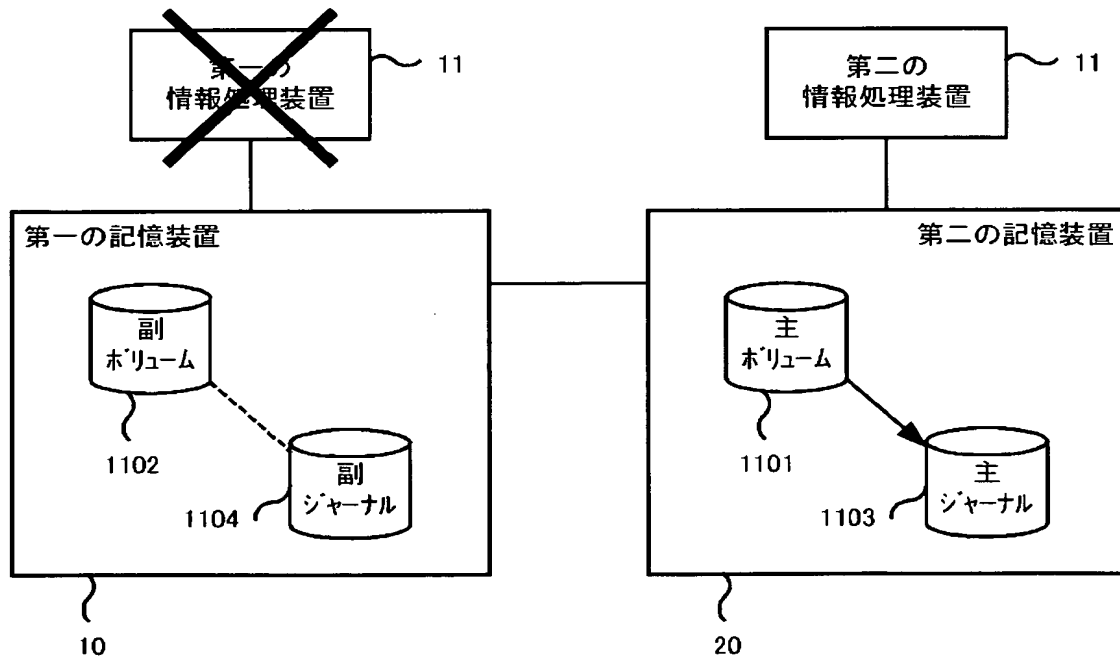
【図 16】



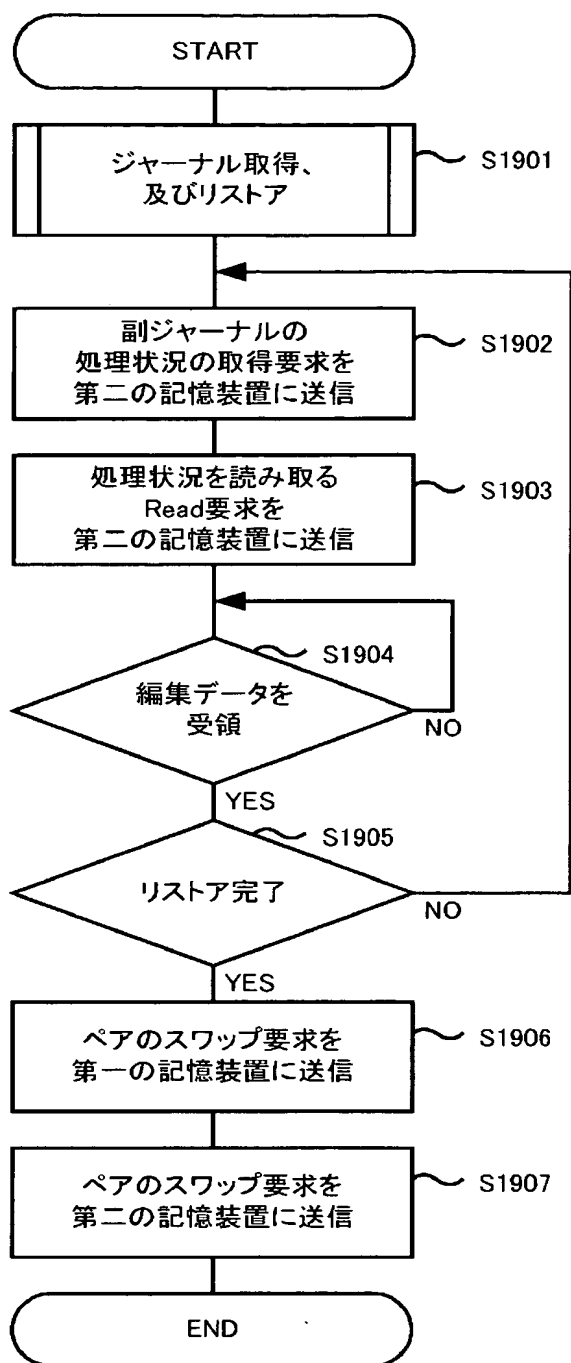
【図 17】



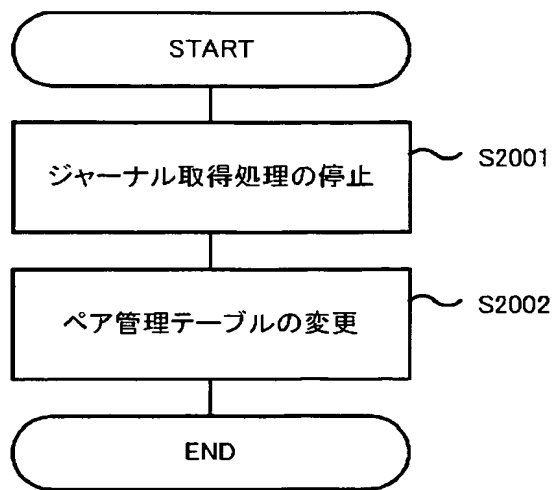
【図 18】



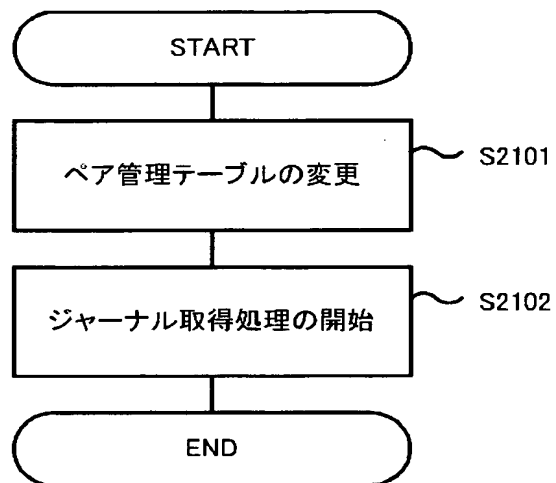
【図 19】



【図 20】



【図 21】



【書類名】要約書

【要約】

【課題】 プライマリサイトとリモートサイトとの間におけるデータのバックアップ運用において、プライマリサイトの情報処理装置からリモートサイトの記憶装置に命令を伝達し、バックアップ運用を制御する。

【解決手段】 情報処理装置が、通常の W r i t e コマンドを用いて、通信可能に接続された第一の記憶装置経由で、第一の記憶装置と通信可能に接続された第二の記憶装置に、制御命令を伝達する。

【選択図】 図 1

特願 2 0 0 3 - 3 2 5 0 8 2

出 願 人 履 歴 情 報

識別番号 [0 0 0 0 0 5 1 0 8]

1. 変更年月日	1 9 9 0 年 8 月 3 1 日
[変更理由]	新規登録
住 所	東京都千代田区神田駿河台 4 丁目 6 番地
氏 名	株式会社日立製作所